Yksilön dataan liittyvät oikeudet Euroopassa ja Yhdysvalloissa

Oulun yliopisto
Tietojenkäsittelytieteiden tutkinto-ohjelma
LuK-tutkielma
Jesse Rontti
15.6.2020

Tiivistelmä

Tässä kandidaatin tutkielmassa analysoitiin yksilön dataan liittyviä oikeuksia, jotka syntyivät Euroopan yleisen tietosuoja-asetuksen, eli GDPR:n, toimesta. Itse GDPR:n käyttöönotto tapahtui vuonna 2018, jonka jälkeen Euroopan jäsenvaltiot ovat tarvittaessa luoneet tai soveltaneet omaa tietosuojalainsäädäntöä GDPR:n säätämien asetusten mukaisesti. Yksilöistä kerätyn datan lisääntyessä yrityksillä ja palveluilla on ollut laajempi pääsy ja saatavuus yksilöiden henkilötietoihin. Tämä lisääntynyt saatavuus johti tutkielman aikaisten ensimmäisten laajempien tietosuoja-asetusten muodostumiseen, joka antoi yksilöille enemmän dataan liittyviä oikeuksia ja loi yrityksistä tietoturvallisempia datan käsittelijöitä.

Tuodessaan yksilöille turvaa, tietosuoja-asetukset asettivat yksilöiden dataa käsitteleville yrityksille rajoituksia, jotka johtivat yrityksien tietosuojakäytäntöjen muutoksiin. Myös reilun datatalouden muodostumiseen tarvittiin lisää käytäntöjä ja ohjausta, joista yksi valmistelluista ja ehdotetuista ratkaisuista on ollut suomalaisen MyData-mallin käyttö. GDPR oli käyttöön tullessaan ensimmäinen laaja tietosuoja-asetus maailmalla, joka oli suunniteltu suojelemaan eurooppalaisia yksilöitä tutkimuksen aikaisen lisääntyneen datan keräämisen aikana.

Muun maailman tietosuojalainsäädännön tilanteen selventämiseksi tutkielmassa käytettiin kirjallisuuskatsausta ja luotiin vertailua, että kuinka laaja ja kattava GDPR:n asettama lainsäädännöllinen viitekehys oli vertaamalla GDPR tietosuoja-asetusta muualla maailmaa sijaitseviin yksilön tietosuojaa koskeviin lakeihin. GDPR:n ja CCPA:n vertailussa molempien tietosuoja-asetuksien todettiin antavan käyttäjille lisää oikeuksia tietojensa hallintaan, mutta GDPR osoittautui laajemmaksi usealla eri osa-alueella, kuten maailmanlaajuisella ulottuvuudellaan.

Asiasanat
GDPR, CCPA, yksilön data, reilu datatalous, data oikeudet

Ohjaaja
Dosentti Raija Halonen

Sisällys

1. Johdanto

Tässä tutkielmassa selvitettiin, mitä oikeuksia yksilöille on muodostunut heidän datansa hallintaan uusien tietosuoja-asetusten käyttöönotossa. Euroopan yleinen tietosuoja-asetus, eli GDPR, otettiin käyttöön vuonna 2018 ja sitä pidetään maailman tiukimpana tietosuojalakina (Wolford, ei päiväystä). Tutkielmassa verrataan Euroopan tietosuoja lainsäädännön tilannetta Yhdysvaltojen tietosuojalakien tutkielman aikaiseen tilanteeseen.

Tutkimuksessa käsiteltiin syitä, jotka johtivat uusien laajojen tietosuoja-asetusten syntymiseen ja sitä, miten uudet tietosuojalait olivat vaikuttaneet tutkimuksen aikana. Vaikka GDPR ei ole ollut olemassa pitkää aikaa tutkimuksen aikana, niin GDPR rikkomuksia silti kertyi yli 300 kappaletta, joiden hallinnollisten sakkojen määrä on noin 450 miljoonaa euroa (GDPR Enforcement tracker, 2020). Tutkimuksessa myös vertailtiin GDPR:n antamia yksilön dataan koskevia oikeuksia Yhdysvaltojen CCPA:n tietosuojalain antamiin oikeuksiin.

Nykymaailmasta kerättävät valtavat määrät dataa ovat muodostuneet ’Big dataksi’, jonka käsittely ei välttämättä onnistu perinteisillä tietokantajärjestelmillä enää, vaan datasta näkyvien ilmiöiden jalostamiseen tarvitaan uusia menetelmiä (Kuo & Kusiak, 2019). Suurten yritysten tietoturvamurrot ovat nostaneet ihmisten tietoisuutta heidän datansa käsittelystä. Lisäksi Euroopan yleisen tietosuoja-asetuksen ollessa pinnalla eurooppalaisten internetin käyttäjien huoli henkilökohtaisen datan väärinkäytöstä on kasvanut. (TNS Opinion & Social, 2017.)

Datan avulla voidaan luoda paljon hyvää yhteiskunnassa, mutta datan käyttäjille pitää laatia selvät pelisäännöt, jotta kuluttajat voivat jatkossakin luovuttaa dataa turvallisemmin ja vähemmän pakotettuina (Sitra, 2019). Datan kasvanutta suosiota kuvaa talousmaailman toteamus, että datan arvo on ylittänyt öljyn arvon (The Economist, 2017). Yritysten välillä tapahtuvan kilpavarustelun datan suhteen on epäilty alkaneen vuonna 2014, jolloin yrityksillä oli mahdollisuus kerätä dataa käyttäjistään, muttei välttämättä vielä tietoa tai taitoa hyödyntää niitä (Laguna, 2014). Laguna (2014) myös pohti, mitä tapahtuisi, jos dataa keräävät jättiläiset Facebook tai Google yrittäisi vaikuttaa ihmisten poliittisiin aatteisiin. Tämä toteutui myöhemmin Cambridge Analytica -nimisen konsultointiyrityksen toimesta Yhdysvaltojen presidenttivaaleissa, jossa käytettiin Facebookin ohjelmointirajapinnan porsaanreikää, joka mahdollisti käyttäjien tietojen saamisen lisäksi tiedonsaannin käyttäjien kavereilta (Isaak & Hanna, 2018).

Voidaan siis pohtia, onko GDPR onnistunut turvaamaan yksilön dataan liittyviä oikeuksia vai onko tavoite reilusta datataloudesta kesken. Reilussa datataloudessa yksilön datan käsittely on läpinäkyvä prosessi, joka ei jätä varaa datan epäeettiseen käyttöön, mutta onko kyseisen reilun datatalouden käsite saavutettavissa vai onko kyseessä vain pelkästään utopistinen haave?

Tämän kandidaatintutkielman tutkimuskysymys oli: Miten voidaan turvata yksilön dataoikeudet tulevaisuudessa?

Koska teknologian edistyessä henkilökohtaisen datan kaupallinen hyödyntäminen on herättänyt uusia huolia liittyen yksilöiden yksityisyyteen, joten tutkielma analysoi, että onko uusien tietosuoja-asetusten antamat oikeudet auttaneet yksilöitä luottamaan paremmin yrityksiin. Tutkielmassa pohdittiin myös, että miten hyvä perusta Euroopalla on reilun datatalouden rakentamiselle. Reilussa datataloudessa yksilön datan käsittely on

läpinäkyvä ja selkeä prosessi yksilölle ja yrityksille, jossa ei jää tilaa epäeettiseen toimintaan (Sitra, 2019) .

Tutkimuksessa analysoitiin mistä yksilöiden data koostuu, mitkä tietosuojalait vaikuttavat yksilön datan käsittely oikeuksiin ja mikä tietosuojalainsäädännön tilanne on Euroopassa ja Yhdysvalloissa. Tiedon haku toteutettiin kirjallisuuskatsauksena eri lähteistä, ja datan merkitystä tutkittiin tieteellisistä näkökulmista ja tietoa tietosuojalainsäädännöstä haettiin hallinnollisista lähteistä.

Tutkielman tarkoituksena oli selvittää, miten yksilöiden dataan liittyvät oikeudet erosivat Euroopassa ja Yhdysvalloissa tutkimuksen aikana. Tutkielmassa pyrittiin myös tuomaan ilmi, miten GDPR on vaikuttanut yritysten tietoturvakäytäntöihin hallinnollisten sakkojen muodossa. Lisäksi tutkielmassa pohditaan reilun datatalouden käsitteen realistisuutta ja tulevaisuutta.

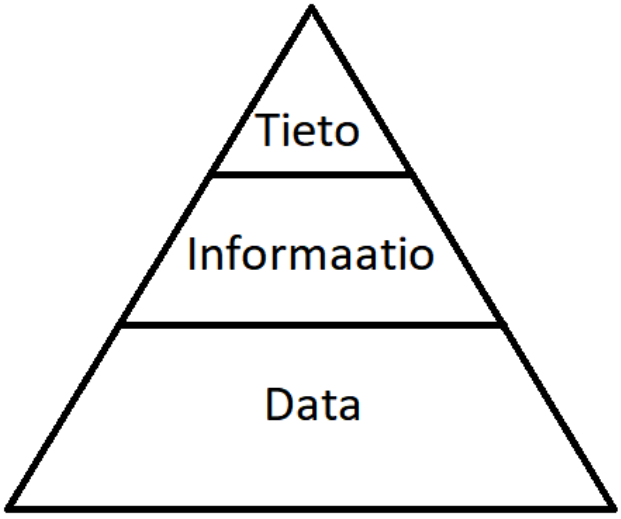
Näiden käsitteiden ymmärtämiseksi ja pohdintaan osallistumiseksi on kuitenkin tärkeä selventää olemassa olevan tutkimuksen perusteella, mitä data todellisuudessa on ja mistä se muodostuu. Dataa käsittelevän luvun jälkeen käydään läpi yksilön tietosuojaa, josta yksilön dataan liittyviä oikeuksia tarkastellaan GDPR:n perusteella. Yksilön tietosuojan läpikäynnin jälkeen yksilön dataan liittyviä oikeuksia verrataan Euroopan ja Yhdysvaltojen välillä. Maanosien välisten tietosuojalainsäädännön vertailun jälkeen tutkimuksen aikana käytyä edellistä tutkimustietoa käydään läpi pohdinnassa ja tutkimuksessa esitetty tutkimuskysymystä reflektoidaan löydösten perusteella.

2. Data

Tässä luvussa tarkastellaan dataa yksilön tasolla ja selvitetään muodostuneen datan alkuperää ja sitä, millä tavoin sitä käsitellään yrityksissä ja palveluissa.

2.1 Datan, informaation ja tiedon määrittely

Yksilön datan arvon ymmärtämiseksi on ensin paneuduttava datan, informaation ja tiedon peruskäsitteisiin. Näiden kolmen käsitteen hierarkkista mallia voidaan kuvata kolmiossa siten, että kolmion pohja taso on itse data, josta pääse keskitasolle informaatioon ja kolmion huippu on itse tieto (Ackoff, 1989). Tämän informaatiomallin hahmotteli ensimmäistä kertaa Ackoff vuonna 1989, kun hän kuvasi tiedon syntymistä. Vaikka Ackoff ei itse artikkelissaan visualisoinut tiedon synnyn prosessia tällä kyseisellä mallinnuksella, niin häntä silti pidetään kyseisen tiedon hierarkia kolmion luojana (Wallace, 2007). Tutkimalla tätä kolmiota voimme ymmärtää, että miten dataa voidaan hyödyntää päätöksenteossa, kuten esimerkiksi markkinoinnissa.



Kuva 1. Tiedon hierarkia kolmio (Wallace, 2007).

Vaikka alkuperäinen Ackoffin artikkeli on vuodelta 1989, niin voidaan sitä silti soveltaa nykyaikaan. Kaikki alkaa kolmion perustasta, eli datasta. Data itsessään on symboleja, joka kertoo datan esittävän aiheen ominaisuuksia. Dataa tuotetaan koko ajan ja kerätään erilaisilla antureilla. Ja nyt esineiden internetin ja sosiaalisten medioiden kukoistaessa kerrotaan, että jopa 90 % maailman datasta on luotu viimeisen kahden vuoden aikana (SINTEF, 2013). Eli dataa on ollut suuria määriä ja sen määrä tulee kasvamaan, kun tilanne on ollut jo kyseinen vuonna 2013. Mutta itse data tällaisenaan ei ole hyödyllistä, vaan se pitää prosessoida ja tarvittaessa jopa puhdistaa, jotta sitä voidaan käyttää hyödyksi (Zhang, Zhang & Yang, 2003).

Datan ja tiedon välissä on informaatio, joka vastaa kysymyksiimme datasta. Informaatio tuottaa vastauksia kysymyksiin liittyen dataan, kuten kenen data on, mitä tämä kyseinen data on, mistä se on saatu ja milloin se on saatu. Informaatio muodostaa suhteen datan ja kysymyksien välille. Näin voimme saada datan kontekstin selville ja tätä kautta tehdä päätöksiä asioista. (Bellinger, Castro & Mills, 2004.)

Tieto on kolmion ylin sarake, joka perustuu tiedetystä informaatiosta. Tieto on tietämyksen ja ymmärryksen sekoitusta, jota voidaan hyödyntää päätöksenteossa (Bellinger et al., 2004). Data-analytiikassa tietoa voidaan soveltaa esimerkiksi hakukoneoptimoinnissa, jossa tekoäly pystyy parantamaan yleisimmin haettujen nettisivujen listausta hakukoneen luettelossa (Pecánek, 2020).

Vaikka Ackoffin teoria tiedon muodostumisesta on vuodelta 1989, niin voidaan sitä silti soveltaa ymmärtääksemme nykyajan data-analytiikkaa. On tärkeää ymmärtää, että data ilman kontekstia ei itsessään tuota yhtään arvoa, vaan meidän pitää jalostaa siitä informaatiota esimerkiksi seuraamalla ilmiöitä, mitä data kertoo. Mallintamalla tietoa sovellamme dataamme ymmärrettävämmäksi, kuten ER-kaavioiden käyttö (eng. *Entity Relation Model*) tietokannassa (Rouse, 2018). Pienemmissä mittasuhteissa voimme taas suosia datan visualisointia, jolloin tarkastelemamme asiat voidaan esitellä esimerkiksi pylväsdiagrammeissa, jolloin erisuurusten määrien vertailu helpottuu ja ilmiöitä on helpompi tunnistaa (Import.io, 2019).

2.2 Yksilön data

Yksilöistä kerätään yhä enemmän tietoja heidän internetin käytön perusteella, mikä itsessään ei ole yllätys ottaen huomioon, että useimmat meistä ovat täyspäiväisesti yhdistettyinä verkkoon älypuhelimillaan. Tietojamme liikutellaan ja yhdistellään sosiaalisista medioista, kyselyistä ja rekistereistä ja näin niitä voidaan hyödyntää moniin eri tarkoituksiin, mutta valitettavasti emme aina ole tietoisia, että mihin. (Vänskä & Härkönen, 2020.)

2.2.1 Henkilötiedot

Henkilötiedoista puhuttaessa tarkoitetaan sellaisia tietoja, joista voidaan tunnistaa yksittäinen henkilö. Näitä ovat esimeriksi nimi, kotiosoite, puhelinnumero, henkilötunnus ja potilastiedot. Näillä tiedoilla ihminen voidaan tunnistaa suoraan, tai näitä tietoja yhdistelemällä henkilö on mahdollista tunnistaa. Tietosuoja-asetus suojaa henkilötietoja riippumatta tiedonkäsittelyn teknologiasta, eli yksilön henkilötietoja tulee suojella, oli tiedot sitten IT-järjestelmässä, perinteisessä paperiarkistossa tai videovalvontajärjestelmässä. (Tietosuojavaltuutetun toimisto, ei päiväystä c.)

Henkilötietoja on mahdollista pseudonymisoida tai anonymisoida. Psuedonymisoidut tiedot lasketaan yhä henkilötiedoiksi, joten näiden käsittelyssä joudutaan soveltamaan tietosuojasäännöksiä. Tämä menettelytapa on tyypillistä tutkimustoiminnassa, jossa henkilöihin voidaan viitata peitenimillä. Anonymisoiduissa tiedoissa ei sovelleta tietosuojasäännöksiä, mutta vaatimuksena tälle on, että kaikki tunnistamiseen liittyvät tiedot on otettu huomioon. (Tietosuojavaltuutetun toimisto, ei päiväystä d.)

Normaalien henkilötietojen lisäksi henkilöllä on erityiset henkilötiedot, joihin kuuluvat henkilön etninen alkuperä, poliittiset mielipiteet, uskonnollinen vakaumus, terveyttä koskevat tiedot, henkilön seksuaalista suuntautumista koskevat tiedot ja geneettistä tai biometristä tunnistamista koskevat tiedot (Tietosuojavaltuutetun toimisto, ei päiväystä e). Näiden erityisten arkaluontoisten henkilötietojen käsittely on lähtökohtaisesti myös kiellettyä, ellei lainsäädännöstä löydy asiaankuuluvaa perustelua EU:n yleisen tietosuoja-asetuksen artikla 9. perusteella (EUR-lex, 2016).

2.2.2 Datan lähteet

Verkkoympäristön bisnesmallina voidaan nykyään pitää käyttäjiin kohdistuvaa tiedonkeruuta. Tämä ilmenee varsinkin nykyisessä verkkosivujen vaihtokaupassa, jossa käyttäjä pääsee käyttämään verkkosivun palveluita vain omat tietonsa luovuttamalla. Tässä vaihtokaupassa ei kuitenkaan ole kysymys meidän henkilötiedoistamme, vaan evästeistä (eng. *Cookies*). Näistä evästeistä muodostuu yksilön data internetissä, jota voidaan pitää digitaalisena jalanjälkenä. (Karhula, ei päiväystä.)

Toinen suuri tiedontuottaja kulkee päivittäin mukanamme joka päivä ja näin voi saada erittäin yksityiskohtaista, ellei jopa arkaluontoisena pidettävää dataa. Puhelimet keräävät passiivisesti paikannustietoja niiden ollessa päällä ja näiden lisäksi muita sovelluksilla tehtyjen toimintojen transaktioista syntynyttä dataa (Binns et al., 2018). Nykyaikana monet viestintäpalvelut ovat yksityisyytemme vuoksi salattuja, mutta käyttäjillä kannattaa silti olla varuillaan, että minkälaisten sovellusten kanssa on valmis jakamaan tietojaan. Tämän tyylisissä datavarannoissa ilman kunnollista suojausta älypuhelimista saatava data voi vaarantaa ihmisten yksityisyyden, vaikka kyseiset tiedot olisivatkin anonymisoitu, niin niistä voisi silti syntyä tunnistuksen uhka (Blondel, 2013).

Kolmas merkittävä yksilön datan lähde on sosiaalinen data. Sosiaalinen data koostuu sosiaalisten medioiden mediasta ja toiminnasta. Maailman laajuisesti sosiaalinen media Twitter on todistanut itsensä hyödylliseksi kehittäjäystävällisellä ohjelmointirajapinnallaan (Bruns & Stieglitz, 2014). Ilmiöiden tunnistaminen on osoittanut helpoksi johtuen Twitterin risuaita (eng. *Hashtag*) -lajitteluteknologiasta. Risuaitojen käytön avulla voidaan helposti rajata julkaisujen aihealue ja tarkastella miten yhteisöt reagoivat eri aiheisiin (Bruns & Stieglitz, 2014).

2.3 Yksilön datan keräämisen tehostuminen

Teknologinen kehitys ja palveluiden digitalisoituminen ovat edesauttaneet kaikkia talouden aloja. Datatalouden arvo ja sen mahdollinen potentiaali on nousussa. Henkilökohtaisen datan on huomattu parantavan palveluja, niin kaupallisia kuin yhteiskunnallisia. Yksilön suostumuksellisen datan jaon kautta on onnistuttu parantamaan yleistä hyvinvointia, talouskasvua ja kustannussäästöjä. Avoimet ohjelmointirajapinnat ovat auttaneet uusien liiketoimintamallien luomisessa ja ovat pienentäneet merkittävästi datan jakamiseen liittyviä kustannuksia. Henkilökohtainen data on mahdollistanut uusia palveluita ja luonut arvoa käyttäjille esimerkiksi hyvinvointisovellusten ja laitteiden muodossa. Tässä on vain osa syistä, mitkä ovat auttaneet ja tehostaneet yksilön datan keräystä ja sen hyödyntämistä. (Sitra, 2019.)

2.3.1 Big data ilmiönä

SAS:in määritelmän mukaan suurta määrää dataa voidaan kutsua Big dataksi, eli massadataksi, kun sen käsittely perinteisillä datan käsittely metodeilla osoittautuu hankalaksi tai jopa mahdottomaksi (SAS, ei päiväystä). Vaikka SQL-peräisten relaatiotietokantojen suosio on ollut vahvaa, niin skaalautuvuus on osoittanut niiden ongelmaksi massadatan prosessoinnissa (Casado & Younas, 2015). Massadatan käsittelyn perinteisillä tietokantajärjestelmillä vaikeaksi tekee jo paikkansa vakiinnuttanut ”**3 V’s of big data**” (Casado & Younas, 2015).

Ensimmäinen tunnusomainen ominaispiirre massadataan liittyen on **määrä** (eng. *Volume*). Määrällä viitataan massadatan suuruuteen ja sen jatkuvaan syntymiseen (Salo, 2013). Suurempi määrä dataa on toki hyödyllistä datan analysoinnissa, mutta haittapuolena vastaavasti suuremmat määrät dataa vaatii enemmän säilytystilaa ja tehokkaampaa prosessointia (Casado & Younas, 2015). Toinen massadatan tunnistettu ominaispiirre on **vauhti** (eng. *Velocity*). Vauhdilla kuvataan kiihtyvää nopeutta, jolla data siirtyy tietokantaan ja siitä käsittelyyn tietojärjestelmissä (Salo, 2013). Massadatan kolmas tunnistettu ominaispiirre on **vaihtelevuus** (eng. *Variety*). Tämä kuvastuu datassa siinä sijaitsevilla epäyhtenäisyyksillä (Salo, 2013). Vaihtelevuutta syntyy, kun dataa kerätään eri formaateista ja näin mediatyypit voivat vaihdella tekstistä videoihin (SAS, ei päiväystä).

2.3.2 Teknologian kehitys

Suurien ja nopeaa kasvavien datamassojen käsittely, säilöntä ja analysointi edellyttävät entistä voimakkaampia työkaluja (Manyika, et al., 2011). Yksi uusista jopa kustannusystävällisistä datan keräys järjestelmistä on uudet avoimet rajapinnat (engl. *Open API*), joita eri järjestelmät, kuten sosiaaliset mediat, tarjoavat. Tämän kautta kehittäjät voivat rakentaa kolmannen osapuolen sovelluksia, kuten esimerkiksi sosiaalisen median Twitterin ohjelmointi rajapinnan kautta, jota tutkijat ovat käyttäneet kerätäkseen lähes reaaliaikaista dataa ihmisten reaktioista nykyisiin tapahtumiin (Bruns & Stieglitz,, 2014). Näin avoimet ohjelmointirajapinnat mahdollistavat suuremman datan saannin, avoimemman datan saannin ja ovat edesauttaneet näin monen uuden palvelun tai tuotteen syntyä.

Datan yhdistely ja säilöntä on enemmän ihmisten saatavilla nykyaikana, johtuen trendeistä, kuten Mooren lain edistämästä kasvaneesta laskenta tehosta ja digitaalisen säilömisen helpottumisesta pilvipalveluiden yleistymisestä (Manyika, et al., 2011). Pilvipalvelut määritellään verkon yli jaettaviksi ja käytettäviksi fyysisiksi tai virtuaalisiksi resursseiksi, joita on mahdollisuus vuokrata kustannustehokkaasti (Murugesan & Bojanova, 2016). Datan käsittelyn helpottumiseen on myös vaikuttanut kehittyneemmät ohjelmistot uusilla edistyneillä teknologioillaan, jotka ovat onnistuneet valjastamaan lisääntyneen laskenta tehon hyötykäyttöön (Manyika, et al., 2011).

Osa nykyisistä tiedonkäsittelyn työkaluista on myös erikoistunut Big datan käsittelyyn, joten vaativan massadatan käsittely onnistuisi paremmin, kuin perinteisillä työkaluilla (Manyika, et al., 2011). Esimerkkinä tähän on Apache Software Foundationin kehittämä Hadoop, joka on suunniteltu suurien tietomassojen hallintaan ja prosessointiin (Apache, ei päiväystä.).

3. Tietosuoja

Tässä luvussa käydään läpi yksityishenkilöiden tietosuojan periaatteita, joissa käy ilmi, miten yksilöillä on oikeus käsitellä tietojaan ja olla tietoisia saamastamme käsittelyistä. Nämä tiedot tulevat ilmi EU:n yleisestä tietosuoja-asetuksesta, jossa kerrotaan yksilön oikeuksista, kun yritys tai organisaatio käsittelee henkilötietoja (EUR-lex, 2016). Luvussa myös käydään läpi, että miten GDPR on vaikuttanut yritysten toimintaan olemassaolonsa aikana hallinnollisten sakkojen muodossa. Luvun lopussa esitellään suomalainen MyData-malli, joka pyrkii ihmiskeskeisellä lähestymistavallaan luomaan reilumman datatalouden mallin, jossa tietojen käyttö on eettisempää ja läpinäkyvämpää.

3.1 Yksilön yleiset data oikeudet

Toukokuussa 2018 vuosien valmistelu saapui päätökseensä ja Euroopan laajuinen yleinen tietosuoja-asetus tuli käytäntöön. Tässä luvussa tarkastellaan yksilön tietojenkäsittelyyn liittyviä oikeuksia, niin kuin ne on ilmoitettu Suomen tietosuojavaltuutetun toimiston nettisivuilla Tietosuoja.fi:ssä, joka toimii Suomen virallisena tietosuojalainsäädäntöä valvovana viranomaisena. (Tietosuojavaltuutetun toimisto, ei päiväystä f.)

Yksilöllä on oikeus saada tieto, että mihin tarkoituksiin hänen henkilötietojansa kerätään ja miten niitä käsitellään. Tämän lisäksi hänellä on oikeus tarkistella, oikaista ja poistaa tietojaan. Nämä tiedot yksilölle tulee tarjoamaan rekisterinpitäjä, eli yritys, viranomainen tai yhteisö, joka vastaa henkilötietojesi käsittelyn tarkoituksista ja keinoista. Yksilö voi myös halutessaan rajoittaa omien tietojensa käsittelyä, vastustaa tietojen käsittelyä ja voi valita olla joutumatta automaattisen päätöksenteon kohteeksi. Kaikkia oikeuksia ei kuitenkaan voida soveltaa esimerkiksi yksilön henkilötietoja ei tarvitse poistaa, jos niiden käyttö on selkeästi perusteltua vaikkapa lain noudattamisen perusteella tai rekisterinpitäjän julkisen vallankäytön perustelulla. (Tietosuojavaltuutetun toimisto, ei päiväystä a.)

Yleisimpiä yksilön tietoa käsitteleviä tahoja, eli rekisterinpitäjiä, ovat työnantajat, pankki ja terveyspalvelut. Kun rekisterinpitäjällä on hallussa yksilön henkilötietoja, on heidän kerrottava yksilölle, että miksi henkilötietoja tarvitaan, kuinka kauan niitä tarvitaan, luovutetaanko henkilötietoja eteenpäin ja valistaa yksilöä, että miten hän voi käyttää henkilötietoihin liittyviä oikeuksia. Nämä tiedot löytyvät yleensä rekisterinpitäjän verkkosivuilta, mutta jos ne jostain syystä puuttuvat, tai yksilö ei löydä niitä, niin tulee tällöin ottaa yhteyttä rekisterinpitäjään. Huomattavaa on myös, että rekisterinpitäjä voi saada henkilötiedot muualta kuin itse yksilöltä, jolloin tiedonkäsittelyyn liittyvät tiedotukset voivat oli monimutkaisempia. (Tietosuojavaltuutetun toimisto, ei päiväystä h.)

Kun yksilö haluaa mahdollisesti tarkastaa tai muokata tietojaan, tulee hänen esittää tarkastuspyyntö rekisterinpitäjille. Tarkastuspyyntö tulee esittää henkilökohtaisesti suoraan rekisterinpitäjälle esimerkiksi sähköpostitse. Tämän jälkeen rekisterinpitäjä on velvoitettu vastaamaan sinulle yhden kuukauden kuluessa, mutta jos pyyntö on monimutkainen tai rekisterinpitäjä ilmoittaa tarvitsevansa lisää käsittelyaikaa, niin muuttuu määräaika tällöin kolmeksi kuukaudeksi. Jos vastausta rekisterinpitäjältä ei kuulu määräajassa, voi yksilö tällöin kääntyä tietosuojavaltuutetun puoleen. Rekisterinpitäjän kieltäytyminen tietojen annosta johtaa sanktioihin tietosuojavaltuutetun toimesta, ellei kieltäytymiseen löydy asianmukaista syytä. Kieltäytyminen tulee perustaa

lakiin, ellei toimijalla ole erityisvaltuuksia, kuten rajoitettu tarkastusoikeus. Rajoitetut tarkastusoikeudet ovat yleensä määritelty yleisen turvallisuuden perusteella ja toimijoille, kuten Poliisi, Rajavartiolaitos, Tulli, Puolustusvoimat ja Rikosseuraamuslaitos. (Tietosuojavaltuutetun toimisto, ei päiväystä b.)

3.2 Euroopan uusi yleinen tietosuoja-asetus GDPR

Euroopan Unioni esitteli tietosuoja-asetuksensa direktiivin, eli kansallisille lainsäätäjille suunnatun toimintaohjeen (GDPR), huhtikuussa 2016. Tietosuoja-asetus on luotu turvaamaan yksityishenkilön henkilötiedot ja näin antaa yksityishenkilölle valtuuksia päättää, että mitä tietoja yritykset omistavat hänestä ja miten hänen tietojansa käsitellään (EUR-lex, 2016). Tämä tarkoitti suurta muutosta ei pelkästään eurooppalaisille yrityksille ja organisaatioille, vaan kaikille toimijoille, jotka käsittelevät Euroopan kansalaisia koskevaa dataa, eli esimerkiksi muiden maanosien yrityksille (Wolford, ei päiväystä). Tämä on johtanut siihen, että yritykset, jotka käsittelevät ihmisten arkaluontoisia henkilötietoja joutuvat lakisääteisesti palkkaamaan tai nimeämään tietosuojavastaavan yritykseensä (Tietosuojavaltuutetun toimisto, 2016). Itse yleistä tietosuoja-asetusta alettiin soveltamaan kaksi vuotta direktiivin antamisen jälkeen toukokuussa 2018 ja tällöin yleinen tietosuoja-asetus kumosi tietosuojadirektiivin 95/46/EY, joka oli annettu vuonna 1995 (EUR-lex, 2016).

GDPR vaatii, että jos henkilötietoja kerätään, niin rekisterinpitäjillä tulisi olla tekniset ja organisatoriset valmiudet käsitellä tietoja, eli tiedolta vaaditaan eheyttä ja luottamuksellisuutta. Suoranaisesti GDPR ei kuitenkaan latele vaatimuksia tieturva käytännöistä ja teknologioista, vaan yrityksen tulee soveltaa tietoturvallisuutta omien periaatteiden mukaan riippuen tiedon arkaluontoisuudesta. Asetuksessa myös korostetaan rekisterienpitäjien vastuullisuutta käsittelytoimista. Tämä tulee ilmi esimerkiksi käyttäjien tekemissä tarkastelupyynnöissä, jolloin yritykset joutuvat tiedottaman käsittelyprosessistaan avoimesti käyttäjälle tietyn määräajan sisällä. (EUR-lex, 2016.)

3.2.1 Suomen tietosuojalaki

Suomessa joulukuussa 2018 julkaistu Tietosuojalaki pyrkii täydentämään Euroopan Unionin yleistä tietosuoja-asetusta ja sen kansallista soveltamista. Laki kertoo tietosuoja-asioita valvovan viranomaisen nimittämisestä, eli tietosuojavaltuutetusta. Lisäksi tietosuojalaissa säädellään lapsiin sovellettavaa ikärajaa tietoyhteiskunnan palveluita tarjottaessa ja erityisten henkilötietoryhmien käsitellystä. (Finlex, 2018.)

Ennen GDPR:n tuloa tietosuoja-asetuksia valvovan viranomaisena on ollut Tietosuojavaltuutettu. GDPR:n tulon myötä kasvaneet tietosuoja-asetuksiin liittyvät asiamäärät ovat kasvaneet ja näin tietosuojavaltuutetuille on myönnetty lisäresursseja. Lisää virkoja on perustettu tietosuojavaltuutetun toimistoon ja aikaisempien virkojen lisäksi viisihenkinen asiantuntijalautakunta avustaa tietosuojavaltuutettua lainsäädännön soveltamiseen liittyvistä asioista. (Oikeusministeriö, 2018.)

Sosiaalisen median ja tietoyhteiskunnan palveluihin sovellettava ikäraja on Suomen tietosuojalain kautta muutettu 16. ikävuodesta 13. ikävuoteen (Finlex, 2018). Tätä nuoremmilla lapsilla on oltava vanhempien suostumus kyseisten palvelujen käyttämiseen (Oikeusministeriö, 2018).

Poikkeuksia Suomen tietosuojalaki kuitenkin esittää sananvapauden, tutkimuksen ja arkistoinnin turvaamiseksi. Eli journalismissa ja tieteellisessä tai historiallisessa tutkimuksessa on mahdollista poiketa tietyistä tietosuoja-asetuksia koskevista velvoitteista, jos tämä on tarpeellista tavoitteiden kannalta. Näiden poikkeuksien tavoitteena on säilyttää nykyisenkaltainen säätely, mutta tärkeää on kuitenkin muistaa, että rekisteröidyn oikeuksien käsittelystä poikkeaminen vaatii aina yksityiskohtaisen vaikutuksenarvioinnin tai eri käytännesäännöt. (Oikeusministeriö, 2018.)

3.2.2 GDPR:n soveltaminen yrityksissä

GDPR:n käyttöönotto on tuonut omat haasteensa yrityksille. IAPP:n tekemän maailmanlaajuisen selvityksen mukaan vuonna 56% yrityksistä ei ole vielä saavuttanut täydellisiä GDPR:n asettamia vaatimuksia tai ei tule noudattamaan asetusta täysin (IAPP, 2018). Yritysten tulee voida selvittää tarvittaessa rekisteröidyn, eli yksilön jolta tieto on kerätty, tekemiä tarkistuspyyntöjä tietoihinsa määräajan kuluessa (Tietosuojavaltuutetun toimisto, ei päiväystä b).

Kun yritys tai yhteisö keräilee ja säilyttää henkilötietoja, niin se on automaattisesti rekisterinpitäjä, joka kantaa vastuun datan säilytyksestä. Rekisterinpitäjän on kartoitettava todennäköiset riskit, jotka saattaisivat kohdistua rekisteröityjen dataan ja tehdä toimenpiteitä varmistaakseen, että asetuksen vaatimuksia tulee noudatettua. Henkilötietojen käsittelijät toimivat nimensä mukaisesti yrityksen henkilötietojen käsittelijöinä ja he toimivat rekisterinpitäjien ohjeiden mukaisesti ja alaisuudessa. (EUR-lex, 2016.)

Rekisterinpitäjien vastuiden lisäksi yrityksen tulee nimetä tietosuojavastaava, joka seuraa asetusten toteutumista henkilötietojen käsittelyssä ja auttaa säännösten tulkinnassa ja noudattamisessa. Tietosuojavastaava on yrityksen sisäinen asiantuntija, joka antaa tietoja tietosuoja-asetuksen mukaisista velvollisuuksista henkilötietoja käsitteleville työntekijöille. Tietosuojavastaava vastaa myös rekisteröityjen yhteyshenkilö tietojen käsittelyyn liittyvissä asioissa, kuten tarkastuspyynnöissä. Hän tarvittaessa on myös yhteydessä ja tekee yhteistyötä valtion virallisen viranomaisen tietosuojavaltuutetun kanssa. Tietosuojavaltuutetulla on myös velvollisuus tiedottaa yksityishenkilöille, jos heidän tietonsa ovat altistuneet tarkastelun kohteeksi tietoturvamurroissa. Tiedotta jättämistä tietoturvamurron tapahtuessa pidetään rangaistavan tekona, josta voidaan jakaa sakkoja riippuen murron laajuudesta. (Tietosuojavaltuutetun toimisto, 2016.)

3.2.3 GDPR rangaistukset

Yleisen tietosuoja-asetuksen noudattamatta jättäminen on vakava rike, joka voi johtaa merkittäviin sakkoihin. Rangaistuksien suuruudessa huomioidaan rikkomuksen luonne, vakavuus ja kesto, sekä rikkomisen tahallisuus tai tuottamuksellisuus. Vastaavasti rikkomuksen sattuessa rangaistusta lieventävinä tekijöinä voi olla rekisterinpitäjän toteuttamat toimet rekisteröidylle aiheutuneen vahingon lieventämiseksi, rekisterinpitäjän mahdolliset aiemmat rikkomukset ja yhteistyö valvontaviranomaisen kanssa haittavaikutusten lieventämiseksi. Hallinnollisen sakon summa voi vaihdella lievemmissä tapauksissa 10 miljoonasta eurosta 2 % maailmanlaajuiseen liikevaihtoon, kun taas vakavammat ja laajemmat rikkeet voivat korottaa hallinnollisen sakon summaksi 20 miljoonaa euroa tai 4 %, riippuen siitä kumpi määristä on suurempi. (EUR-lex, 2016.)

Kuten hallinnollisista sakoista voidaan huomata, niin yleisen tietoturvatason ja rekisteröityjen tietojen käsittelyn suojaamista on varmennettava, eli käyttäjien tietosuoja on otettava tulevaisuudessa huomioon ansaitsemallaan tavalla, jotta yritykset eivät joudu sakotuksen kohteeksi. GDPR sakkojen määrää tarkkaileva sivu *enforcementtracker.com* kertoo meille, että tällä hetkellä Euroopassa on jaettu 282 rangaistusta GDPR rikkomuksia koskien (GDPR Enforcement tracker, 2020).

Toukokuun lopussa vuonna 2020 jaettiin ensimmäiset GDPR-sakot Suomessa tietosuojavaltuutetun toimesta. Suurimmat rangaistukset Suomessa on saanut Posti, jonka hallinnollinen sakko oli 100 000 euroa. Syynä rangaistukseen oli Postin tietojenkäsittely toiminnan avoimuuden puute, koska Posti ei ollut kertonut rekisteröidyille heidän oikeuksistaan kieltää tietojen luovuttamista muutosilmoitusten tekemisen yhteydessä. Tietosuojavaltuutettu myös tiedottaa, että Posti oli tarjonnut kyseistä mahdollisuutta vain asiakkaille, jotka olivat ostaneet lisäpalveluita muuttoilmoitusta tehdessä. Asia tuli ilmi tietosuojavaltuutetulle kannelleilta henkilöiltä, joille oli tullut yhteydenottoja ja suoramarkkinointia eri yrityksiltä muuttoilmoituksen tekemisen yhteydessä. Tietosuojavaltuutetun toimiston selvityksessä kävi lopulta ilmi, että posti ei ollut kertonut rekisteröidyille mahdollisuudesta kieltäytyä tietojen luovuttamisesta ja näin 161 000 asiakasta altistui rikkomuksella jo pelkästään vuoden 2019 aikana. (Tietosuojavaltuutetun toimisto, 2020.)

3.3 Miksi meidän tulee välittää data oikeuksistamme?

Kaikista hyödyistä huolimatta datatalous herättää epäilyksiä eurooppalaisten keskuudessa. Tutkimus osoittaa, että 45% eurooppalaisista internetin käyttäjistä kokevat henkilökohtaisen datan mahdollisen väärinkäytön verkkopalveluiden käytön estävän tekijänä vuonna 2017 (TNS Opinion & Social). Tämä on myös saanut käyttäjät ryhtymään toimiin kyseistä ilmiötä vastaan, koska kyseinen tutkimus myös kertoo, että 39% eurooppalaisista internetin käyttäjistä on rajoittanut heidän jakamaansa henkilökohtaisen datan määrää (TNS Opinion & Social, 2017). Kyseiset luvut viestivät kasvaneesta ymmärryksestä ilmiötä kohtaan varsinkin, kun kyseinen tutkimus tapahtui ennen korkean profiilin omaavaa Cambridge Analytican (CA) tapausta, jossa paljastui henkilökohtaisen datan väärinkäytön uhat suuremmalle yleisölle.

Näiden skandaalien tarkastelu auttaa myös ymmärtämään, että miksi ja miten meillä on mahdollisuus nyt rakentaa reilu datatalous uusien tietosuoja-asetusten avulla ja kuinka meillä on mahdollisuus kitkeä harmaita alueita, joita on onnistuttu menneisyydessä epäeettisesti hyödyntämään.

3.3.1 Cambridge Analytica -tapaus

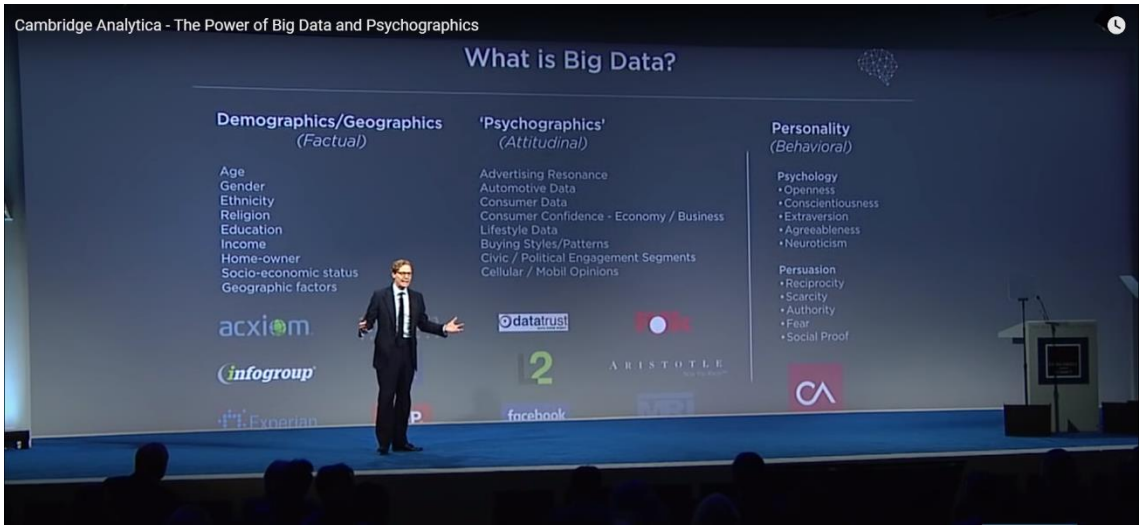
Vuonna 2013 Cambridgen yliopiston psykometriikan tutkijat loivat persoonallisuus testin Facebookiin, jossa käyttäjät pystyivät mittaamaan ”OCEAN” (*openness, conscientiousness, extraversion, agreeableness, and neuroticism*) psykologisen profiilinsa. Tutkimus saavutti 350 000 yhdysvaltalaista osallistujaa ja onnistui luomaan selkeän yhteyden käyttäjän Facebook- tykkäysten ja käyttäjän OCEAN profiilin välillä (Isaak & Hanna, 2018). Eli ihmisiä, heidän persooniaan ja käyttäytymistapojaan pystyttiin mittaamaan Facebook- tykkäysten avulla ilman sen kummempia suorituksia, kuin yksi persoonallisuus testi Facebookissa. Tämä oli edellytys Cambridge Analytican toiminnalle, jonka toiminta hyödynsi Michal Kosinskin tutkimusta, jossa profiloitiin internetin käyttäjiä netistä löytyvällä datalla, Facebook tykkäyksillä ja älypuhelin datalla (Kosinski, Stillwell & Graepel, 2013). Kuvassa 2. CA:n toimitusjohtaja Alexander Nix avaa OCEAN- persoonallisuustestin sisältöä.

Kuva 2. Cambridge Analytican konferenssiesitys (Youtube, 2016)[1].

Iso-Britannialainen poliittinen konsulttiyritys CA näki kyseisen tutkimuksen hyödyn ja onnistui integroimaan sen hyödyt oman alansa käytäntöihin. Vuonna 2014 CA kehitteli oman Facebook- sovelluksensa yhdessä Alexander Koganin kanssa, joka oli tutkija Cambridgen yliopistossa OCEAN persoonallisuustestin kehityksen aikana. Sovellus käytti Facebookin avointa ohjelmointirajapintaa ja onnistui saavuttamaan 87 miljoonan käyttäjän tiedot laittomasti käyttämällä Facebookin sen aikaista porsaanreikää, joka mahdollisti tietojen saannin CA:lle ei pelkästään käyttäjiltä, vaan myös käyttäjän kaikilta kavereilta. Yhdistelemällä dataa sosiaalisesta mediasta, selaimista, internet ostoskäyttäytymisestä ja heidän kehittämästä persoonallisuustestistä CA pystyi analysoimaan yli 5 000 tietopistettä noin 230 miljoonalta amerikkalaiselta äänestäjältä. (Isaak & Hanna, 2018.)

[1] Youtube.com. (2016). Cambridge Analytica - The Power of Big Data and Psychographics. Lainattu 6.7.2020, saatavilla: https://www.youtube.com/watch?v=n8Dd5aVXLCc

CA:n toiminnan metodologia keskittyi kolmeen pääpisteeseen. Käytöspsykologia on ensimmäinen pääpiste, jonka avulla CA kehitti OCEAN- persoonallisuustestin ja näin he pystyivät tunnistamaan, että miten he voivat saada äänestäjien käytöksen muuttumaan, eli miten he voivat saada henkilön äänestämään CA:n edustamaa poliitikko kandidaattia. Toinen pääpiste on Big data, jonka avulla CA pystyi luokittelemaan valtavia määriä dataa äänestäjiltä heidän internet käyttäytymisen perusteella. Kuvassa 3. Alexander Nix esittelee Big dataan sisältyviä tietoja yksilöistä. Kolmas pääpiste CA:n metodologiassa on kohdistettu mainonta, jolla voidaan saada yrittää äänestäjä vaihtamaan poliittista mielipidettään suosien CA:n edustama kandidaattia. Nämä kaksi ensimmäistä pääpistettä oli tärkeässä asemassa, jotta CA tietäisi, että keneen mainonta kannatti kohdistaa ja mihin yksilön periaatteisin kannattaisi vedota. (Vice, 2017.)



Kuva 3. Alexander Nix esittelee CA:n keräämiä tietoja (Youtube, 2016)[2].

Persoonallisuuden määrittelyn avulla CA onnistui lokeroimaan datastaan äänestäjät, jotka eivät olleet vielä varmoja äänestysvalinnastaan, eli käännytettävät (eng. *Persuadables*). Tällöin kohdennettu mainonta astui peliin, jossa pyrittiin käännyttämään äänestäjä äänestämään CA:n ehdokasta, ja yleisinä käytäntöinä mainoksissa oli, että CA:n ehdokkaan ja äänestäjän samankaltaisuuksia suosittiin ja vastaehdokkaita pyrittiin mustamaalaamaan tylysti. (Vice, 2017.)

CA:n epäeettinen elinkaari päättyi vuonna 2018, kun Facebookin tietovuodot tulivat julki, mutta CA kerkesi vaikuttaa Yhdysvaltain vuoden 2016 presidentin vaaleissa (Vice, 2017). CA säilytti Facebookista saamansa dataa Yhdysvaltain vaalien ajan, vaikka he olivat saaneet käskyn tuhota kyseinen data jo vuonna 2015 (Dwoskin, 2018).

[2] Youtube.com. (2016). Cambridge Analytica - The Power of Big Data and Psychographics. Lainattu 6.7.2020, saatavilla: https://www.youtube.com/watch?v=n8Dd5aVXLCc

3.3.2 Data oikeudet suojelevat yksilöitä

Sosiaalisen median merkitys on kasvanut yksilön elämässä, kun yksilöt kommunikoivat entistä enemmän sosiaalisessa mediassa (Hussey, 2020). Tämän kommunikaation yhteydessä syntyy paljon henkilökohtaista dataa yksilön taustasta ja henkilökohtaisista suhteista eri ihmisiin ja yrityksiin erilaisten media muotojen yhteydessä (Lamdan, 2015). Lamdan (2015) kertoo sosiaalisesta mediasta kerätyn datan muodostuvan keskusteluhistoriasta, päivityksistä, käyttäjän julkaisemasta mediasta, GPS-sijainneista, käyttäjän tykkäämistä sivuista, kirjautumisen ajankohdista ja jopa klikkauksista, joita käyttäjä tekee sivulla. Husseyn (2020) mukaan tällä hetkellä sosiaalisen median keräämä datan määrä on ylenpalttista ja sen hyödyntäminen vain kaupallisiin tarkoituksiin tulisi olla rikollista.

Data ja yksityisyys oikeudet suojaavat yksilöitä väärinkäytöltä ja tukevat yksilön perusoikeuksia, koska kaikilla ihmisillä on oikeus suojella yksityiselämäänsä (Finlex, 1999). Oikeaoppinen ja läpinäkyvä tieto henkilökohtaisen datan käsittelystä kasvattaa luottamusta yksilön ja palveluiden välillä (Sitra, 2019). Yksilöllä on suurempi luottamus henkilökohtaisten tietojen käsittelyä kohtaan, kun heillä on tietoa henkilötietojen käsittelyn yksityiskohdista ja tieto siitä, että jos heille suotuja oikeuksia rikotaan, niin rikkomuksilla tulee olemaan seurauksia (Soken-Huberty, ei päiväystä). GDPR on mahdollistanut yksilöille oikeuden pitää yllä mainettaan oikeudella tulla unohdetuksi, jonka avulla yksilöt voivat säädellä, että mitä tietoa hänestä näkyy julkisessa internetissä (Soken-Huberty, ei päiväystä).

3.4 Suomen osuus reilun datatalouden rakentajana

Suomi on aktiivisesti pyrkinyt digitaalisten henkilötietojen käsittelyn edelläkävijäksi esimerkiksi MyData 2018 konferenssilla, johon osallistui poikkitieteellisesti kouluttautuneita huippuasiantuntijoita (Uusiteknologia.fi, 2018). MyData-mallissa on kyse henkilötietojen hyödyntämismallista, jossa yksilöllä on pääsy ja kontrolli häntä koskeviin tietoihin (Liikenne- ja viestintäministeriö, 2018).

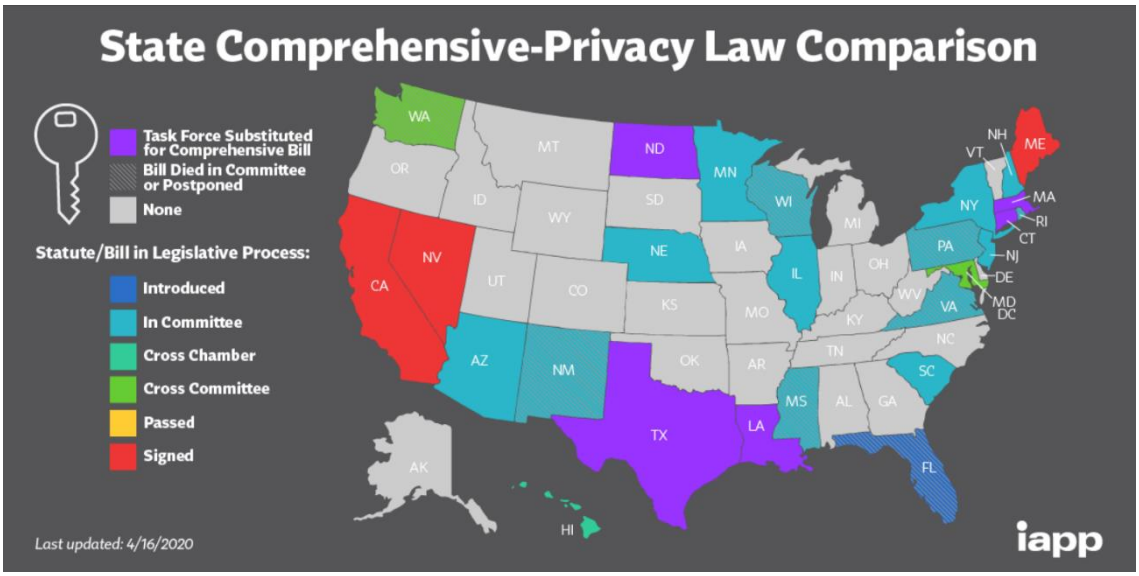
Avoimen omadata-mallin palveluinfrastruktuuria kuvaillessa mallin mainitaan kuvaavan tapaa, jolla rahavirtojen liikkumista hallitaan nykypäivinä, eli monet käyttävät samaa rahaa, mutta toimijoiden välillä on yhteistoiminnan periaatteet ja näin esimerkiksi pankin vaihtaminen onnistuu helposti. Tällaisessa tapauksessa pankin vaihtamisella kuvataan tiedon siirtoa eri rekisterinpitäjien välillä, eli rahan tapaan, tieto ei häviä siirtyessään. Tällainen toteutus voisi luoda yritykselle käyttöönottaessa vaikeuksia saada toiminta etenemään, mutta toiminnassa se edistäisi helppokäyttöisyyttä ja vahvistaisi datan ingretiteettiä. (Poikola, Kuikkaniemi, Kuittinen, O., 2018.)

MyData- ajattelun kolmeksi lähtökohdaksi luonnehditaan ihmiskeskeisyys, tiedon hyödynnettävyys ja liiketoiminnanmallien avautumista (Poikola et al., 2018). Poikola et al., (2018) kertoo näiden periaatteiden ohjaavan MyData- rajapintojen, standardien, palveluinfrastruktuurin, sovellusten ja palveluiden kehitystä. MyData on pyrkinyt pääsemään irti datan omistajuus käsitteestä, koska tiedon suhteen omistaminen ei ole yhtä suoraviivaista, kuin irtaimiston tai kiinteän ominaisuuden omistaminen, joten omistajuuden sijaan MyData pyrkii puhumaan hallinnasta (Pitkänen, 2014). Käytännössä MyData-malli tarkoittaa mahdollisuutta siirtää palvelusta toiseen ensisijaisessa palvelussa kertyneet henkilötiedot, ostoshistoria, kulutusmieltymykset ja rekisteritiedot (Poikola et al., 2018).

Maailman laajuisesti Mydata-malli on kerännyt suosiota järjestettävillä konferensseilla, jonne on kutsuttu asiantuntijoita useilta eri aloilta. Itse MyData Global järjestö on perustettu vuonna 2018 ja siihen on liittynyt yli 80 eri organisaatiota (Mydata.org, ei päiväystä). Suomessa Mydata-mallia on kokeiltu Traficomin kanssa ajo-oikeuksien suostumuspohjaisessa hyödyntämisessä, Oriolan terveystutkimuksessa ja TEM:n yhteistyöhankkeessa työllistymisen ja rekrytoinnin parissa (Vastuugroup, ei päiväystä).

4. Yksilön data oikeudet Euroopassa ja Yhdysvalloissa

Tässä luvussa käsitellään tapauskohtaisia käyttötapauksia sekä Euroopan että Yhdysvaltojen lainsäädännön näkökulmasta. Euroopan näkökulmasta sovelletaan tapauksia GDPR:n mukaan ja Yhdysvaltain lainsäädännön mukaan sovelletaan CCPA:ta (*California Consumer Privacy Act)*, jota voidaan pitää Yhdysvaltain laajimpana tietosuojalakina. Huomautettavaa kuitenkin on, että tämä kyseinen laki koskee vain Kalifornia osavaltion asukkaita, joten tilanne voi olla, että joillakin osavaltioilla ei ole ollenkaan säännöstelty yksilöiden tietosuojaa. Kuvassa 4. voimme nähdä, että millä osavaltioilla on olemassa oleva tietosuojalaki ja, että onko niillä mahdollisesti vireillä uuden tietosuojalain säätäminen. Kuvasta myös paljastuu, että millä osavaltioilla ei ole olemassa olevaa tietosuojalakia.



Kuva 4. Yhdysvaltojen tämänhetkinen tietosuoja lakien tilanne (Iapp.com, 2020)[3]

4.1 Tietosuoja-asetukset Euroopassa ja Yhdysvalloissa

GDPR on edistynein ja laajin yksilön data oikeuksia koskeva lainsäädännön viitekehys maailmassa. Se suojelee Euroopan kansalaisia ja heidän koskemaa dataansa jopa EU-alueiden ulkopuolella, jos yritys käsittelee EU- jäsenten dataa. Tietosuoja-asetuksen säännökset ovat lainsäädännöllisiä, joten niiden käyttöönottoa valvotaan tarkasti. (Wolford, ei päiväystä.)

Yhdysvalloissa tilanne ei ole näin yksinkertainen eriävien osavaltio lainsäädäntöjen takia. Tietosuoja ei ole niin laajasti valvottua yleisellä tasolla, kuin Euroopassa, vaan säädetty osavaltio kohtaisesti. Tämä tarkoittaa, että Yhdysvalloissa yksilön tiedon etsintä on

[3] Iapp. (2020). US State Comprehensive Privacy Law Comparison. Lainattu 13.6.2020, saatavilla: https://iapp.org/media/pdf/resource_center/State_Comp_Privacy_Law_Map.pdf

yleistä kolmannen osapuolen luottoraporteista työnhaussa, terveydenhuollossa tai lainahakemuksissa. (Noordyke, 2020)

4.1.1 GDPR:n ja CCPA:n vertailu

GDPR:ssä kaikilla eurooppalaisilla on oikeus omaan dataansa. Luvussa 3.2 tarkastelimme GDPR:n pääpiirteitä ja pystyimme huomaamaan, että laki kohtelee kaikkia EU:n jäseniä tasapuolisesti käyttöoikeuksien kannalta, eli yksilön ei tarvitse olla yrityksen asiakas voidakseen tarkastella tietoja tai muuttaa niitä, Eli yksilöistä, joista dataa kerätään omaavat automaattisesti ihmisoikeuksien tyylisesti GDPR:n säätämät dataoikeudet. GDPR:ssä rekisteröidyn henkilötietoihin määritellään kaikki tiedot, jotka voivat johtaa luonnollisen henkilön tunnistamiseen tai epäsuoraan tunnistamiseen. Ja henkilöitä, joilta tieto on kerätty, kohdellaan rekisteröityinä (eng. *data subject*). Rekisteröityjen henkilötietoja prosessoivia yrityksiä tai palveluita kohdellaan rekisterinpitäjinä (eng. *data controller*). Tähän lukeutuu kaikki yritykset ja palvelut ilman kokorajoituksia, jotka käsittelevät rekisteröityjen dataa. (EUR-lex, 2016.)

CCPA:ssa Kalifornialaisilla kuluttajilla on oikeus tarkastella omaa dataansa. Kalifornian yksityisyyden suojaa koskeva laki, eli CCPA, on Kalifornian osavaltion tietosuoja-asetus, jonka avulla Kaliforniassa asuvat kansalaiset voivat tarkastella heistä käsiteltäviä tietoja ja mahdollisesti myös kieltää heidän tietojensa eteenpäin myynnin. CCPA:ta voi soveltaa, kun Kalifornian kansalainen on jonkun yrityksen asiakas ja hänestä on kerätty tietoja. Eli yksityisyyden suoja ei ole automaattisesti Kalifornian kansalaisilla, vaan heidän täytyy olla tietoa keräävän yrityksen asiakas, eli kuluttaja (eng. *consumer*). Kuluttajan henkilötiedoiksi luetaan CCPA:n mukaan kaikki informaatio, joka kuuluu tunnistettavalle luonnolliselle henkilölle tai hänen taloudelleen. CCPA ei myöskään käsittele yrityksiä ja palveluita, jotka keräävät tietoa, rekisterinpitäjinä, vaan heitä käsitellään liiketoimina (eng. *business*). Eroten GDPR:n määrittelystä, liiketoimilla on rajoituksia, että milloin he ovat velvollisia noudattamaan CCPA:ta. Liiketoimi on velvollinen soveltamaan CCPA:ta jos yksi näistä ehdoista täytyy; kun liiketoiminnan vuotuinen liikevaihto on yli 25 miljoonaa dollaria, jos liiketoiminta käsittelee ainakin 50 000 Kalifornialaisen henkilötietoja vuosittain tai jos liiketoimen liikevaihdosta muodostuu yli 50 % henkilötietojen myynnistä. (CCPA, 2018.)

4.1.2 Oikeuksien hyödyntäminen

On tärkeää muistaa, että GDPR ja CCPA määrittelee dataa keräävät yritykset ja palvelut eri nimikkein ja myös henkilöt, joista data on kerätty. GDPR:n datan kerääjistä käytetään niistä määriteltyä termiä, eli rekisterinpitäjää, kun taas CCPA:ta soveltaessa viitataan tiedonkerääjiin liiketoimina (EUR-lex, 2016; CCPA, 2018). Myös käyttäjät, joilta tieto kerätään, määritellään erillä tavoin, kuten 4.1.1 luvussa selvisi, mutta heihin viitataan termillä yksilö tässä luvussa selventämiseksi. Eri oikeuksien vertailun apuna on käytetty Dataguidance:n ja Future of Privacy Forumin tekemää julkaisua ” Comparing privacy laws: GDPR v. CCPA”, joka pyrkii vertailemaan kyseisiä lainsäädäntöjä kohta kohdalta (Future of Privacy forum, ei päiväystä).

Molemmat GDPR ja CCPA määrittelevät yksilölle oikeuden poistaa heistä kerättyä dataa. GDPR:ssä tämä oikeus koskee kaikkea dataa muutamia poikkeuksia lukuun ottamatta, mutta CCPA:ssa määrittely koskee vain dataa, jonka liiketoimi on kerännyt käyttäjältä,

eli Kalifornialaisten oikeus tulla unohdetuksi ei toteudu täysin, kun he vaativat henkilötietojen poistoa liiketoimilta. Molempiin lainsäädäntöihin pätee poikkeustilat tiedon poistoihin liittyen, jos tieto on lainsäädännöllistä eikä sitä voida poistaa tai jos tietoa käytetään tutkimukseen tai arkistointiin. (Future of Privacy forum, ei päiväystä.)

Oikeutta saada tietoa henkilötietojen käsittelyssä harjoitetaan molemmissa lainsäädännöissä, eli molemmissa lainsäädännöissä yksilöille tiedotetaan dataa kerättäessä, että mihin tietoa tullaan käyttämään. GDPR painottaa, että rekisterinpitäjien on annettava selkeä huomautus yksilöille, jos heidän tietojaan aiotaan jakaa kolmansille osapuolille, kuten vaikka suoramarkkinointiin. CCPA:ssa tällainen tieto ei ole oletuksena, vaan jos yksilö ei halua, että liiketoimi jakaa hänen tietojaan muille osapuolille voi yksilö tällöin jättäytyä pois kyseisistä toimista, eli estää hänen tietojensa jakamisen kolmansille osapuolille. (Future of Privacy forum, ei päiväystä.)

GDPR ja CCPA tarjoaa yksilöille mahdollisuuden vastustaa heidän henkilötietojen jakamisen ja myymisen kolmansille osapuolille. GDPR:ssä oikeus vastustaa henkilötietojenkäsittelyä on laajempi, koska se koskee henkilötietojen jakamista, prosessointia ja myyntiä, kun taas CCPA:n oikeus vastustaa henkilötietojen jakamista koskee vain myyntiä. GDPR:n avulla yksilöillä on myös oikeus vastustaa suora markkinointia heidän tietojensa perusteella. (Future of Privacy forum, ei päiväystä.)

Yksilön oikeus käsitellä ja päästä käsiksi omiin henkilötietoihinsa auttaa yksilön ja tiedonkerääjän luottamussuhteessa, kun toiminta muuttuu läpinäkyvämmäksi ja avoimemmaksi. Molemmissa lainsäädännöissä tämä toteutuu ja yksilö voi tarkastuspyynnöllä päästä käsiksi hänestä kerättyihin henkilötietoihin. (Future of Privacy forum, ei päiväystä.)

4.1.3 Ulottuvuus

Myös GDPR:n ja CCPA:n lainsäännölliset ulottuvuudet rajoittuvat eri tasoille. Yksilön tasolla GDPR suojelee rekisteröityjä, eli luonnollisia henkilöitä, joille ei ole erikseen määritelty kotipaikkakuntaa tai kansalaisuuksia. CCPA taas suojelee kuluttajia, jotka ovat luonnollisia henkilöitä ja asuvat vakituisesti Kaliforniassa. Näiden kuluttajien tulee myös olla asiakkaita liiketoimille, joista he aikovat saada selviteltyä henkilötietojaan. (Future of Privacy forum, ei päiväystä.)

GDPR myös vaikuttaa kaikkiin yrityksiin, jotka tarjoavat palveluitaan ihmisille Euroopan sisällä, vaikka yrityksellä ei välttämättä olisi itse toimintaa Euroopassa. Näin GDPR vaikuttaa myös muihin maanosiin, kuin pelkästään Eurooppaan ja yritysten on pitänyt päivittää tietosuojaustaan GDPR:n mukaan, jos he ovat halunneet jatkaa työskentelyä eurooppalaisten kanssa. CCPA taas vaikuttaa yritysten toimintaan keillä on toimintaa Kaliforniassa. Tämä määrittely on epäselvä, koska yrityksiin voi lukeutua organisaatiot, nettisivut ja eri palvelut. CCPA:n tarkistuspyynnöissä katsotaan tapauskohtaisesti, että voidaanko CCPA:ta soveltaa. Mutta pääosin Kalifornialaiset voivat hyödyntää oikeuksiaan yrityksiin, joilla täyttyvät 4.1.1 luvussa mainitut koko vaatimukset, jotka keräävät henkilötietoja, päättää henkilötietojen keräämisestä ja harjoittavat liiketoimintaa Kaliforniassa. (Future of Privacy forum, ei päiväystä.)

4.1.4 Lain valvonta

Euroopassa GDPR:ää valvontaan on osoitettu joka maassa oma tietosuojavaltuutettu, joka on valtion määräämä virallinen taho, ketkä valvovat yksilöiden tietosuojaoikeuksia. Tietosuojaviranomaiset tekevät selvityksiä ja tarkastuksia ja määrää hallinnollisia seuraamuksia tietosuoja-asetusten rikkomisesta tarvittaessa (EUR-lex, 2016). CCPA:n tietosuojavaltuutettuna toimii Kalifornian yleinen syyttäjä (eng. *Attorney general*) (CCPA, 2018).

Molemmat lainsäädännöt käyttävät hallinnollisia sakkoja, jos liiketoimet tai rekisterinpitäjät osoittautuvat tottelemattomiksi (Future of Privacy forum, ei päiväystä). GDPR sakotuksen hoitaa tietosuojavaltuutettu ja sakkojen summa voi olla lievissä tapauksissa maksimissaan 2% maailmanlaajuisesta liikevaihdosta tai 10 miljoonaa euroa, kun taas vakavissa rikkomuksissa hallinnollisen sakon summa voi kasvaa 4 % maailmanlaajuisesta liikevaihdosta tai 20 miljoonaa euroa, riippuen siitä kumpi on korkeampi. Rikkomusten tapahtuessa rikkomuksen luonne voi vaikuttaa paljon hallinnollisen sakon summaan, kun yritys voi rikkomusten sattuessa lieventää mahdollista hallinnollisen sakon summaa osoittamalla yhteistyöhalukkuutta ongelman korjaamista kohtaan (Tietosuojavaltuutetun toimisto, ei päiväystä g). Kalifornian yksityisyyttä suojaavassa laissa hallinnolliset sakot ovat lievempiä, kun niiden summa on 2 500 dollaria per rikkomus, tai 7 500 dollari per tahallinen (CCPA, 2018). CCPA:ssa (2018) ei ole mainittu maksimi sakko summaa. Molemmat GDPR sekä CCPA tuomitsevat tietoturvamurrot ja niitä valvovat viranomaiset saattavat voida sakottaa liiketoimia ja rekisterinpitäjiä, jos yksilöiden henkilötiedot ovat vaarantuneet ja heille ei tiedoteta asiasta määrätyn ajan sisällä (EUR-lex, 2016; CCPA, 2018).

4.2 Vertailun yhteenveto

Kalifornian yksityisyyttä suojaava laki, eli CCPA, oli ensimmäinen laajempi kuluttajan yksityisyyttä suojaava tietosuojalaki Yhdysvalloissa, kun se otettiin käytäntöön vuoden 2020 alussa (Jehl & Friel, 2018). Kuvaa 4. tarkastelemalla voidaan kuitenkin todeta, että maanlaajuisesti tilanne ei ole näin hyvä kuin mitä Kalifornian osavaltion kansalaisilla on. Kalifornia on kuitenkin tunnettu Yhdysvalloissa ”trendien asettajana” uusien lainsäädännön käyttöönotoissa, joten voi hyvinkin olla mahdollista, että tulevaisuudessa näemmä vastaavien tietosuojalakien käytäntöjen yleistymisen koko maan tasolla (Paycom, 2018).

Itse CCPA:n ja GDPR:n vertailussa voimme huomata paljon samankaltaisuuksia, mutta suurin ero sijaitsee asetusten ulottuvuudessa ja yksilön määrittelyssä keneltä data kerätään. CCPA:n valvonta perustuu Kalifornialaisten kuluttajien auttamiseen, kun taas GDPR takaa kaikille yksilöille Euroopan alueella dataan liittyviä oikeuksia. Myös valvottavien kohteiden, eli liiketoimien (CCPA) ja rekisterinpitäjien (GDPR), määrittely eroaa toisistaan. CCPA valvoo yrityksiä, joilta täyttyvät tietyt kokoon tai liiketoimintaan liittyvät periaatteet: 1) jos yrityksen kokonaisliikevaihto on yli 25 miljoonaa dollaria, 2) jos yritys kerää yli 50 000 kuluttajan tietoja vuodessa tai 3) jos yrityksen 50% tai enemmän vuosittaisesta liikevaihdosta syntyy kuluttajien datan myymisestä. Kun taas GDPR koskee kaikkia palveluita, yrityksiä ja organisaatioita ympäri maailmaa, jotka pitävät hallussaan eurooppalaisten henkilötietoja. (Jehl & Friel, 2018.)

Oikeuksia tarkastellessa CCPA ja GDPR perustaa toimintansa samalla tavalla yksilön dataa suojaaviin oikeuksiin, kuten oikeudet tarkastella, muokata ja poistaa yksilöä koskevia henkilötietoja. GDPR tosin tarjoaa laajemman oikeuden tulla unohdetuksi

kokonaan rekisterinpitäjiltä kuin CCPA. GDPR myös antaa yksilölle oikeuden olla tulematta automaattisen käsittelyn kohteeksi, jota CCPA ei tue. Kokonaisvaltaisesti GDPR tarjoaa enemmän oikeuksia yksilölle koskien oikeutta vastustaa yksilön tietojen käsittelyä. CCPA tarjoaa vaan mahdollisuuden jättäytyä pois tietojen myynnistä, kun GDPR:n avulla yksilö voi vastustaa profiloinnin, suoramarkkinoinnin tai tutkimuksen tekemisen kohteeksi tulemista. Molemmissa tietosuoja-asetuksissa voidaan käyttää hallinnollisia sakkoja rangaistuksen keinoina ja GDPR on niiden osalta rankempi. Mutta molempien asetusten hallinnollisten sakkojen summat voi mahdollisesti aiheuttaa yrityksillä taloudellista epävakautta. (Jehl & Friel, 2018.)

5. Pohdinta

Kaiken kaikkiaan EU:n yleinen tietosuoja-asetus, eli GDPR, on osoittanut hyvää esimerkkiä maailmalla yksilön tietosuojan ja dataan liittyvien oikeuksien turvaamisessa. Se on onnistunut luomaan viitekehyksen lainsäädännölle, jota EU:n jäsenmaat pystyvät soveltamaan ja tarvittaessa soveltamaan omalla lainsäädännöllä. Vaikka GDPR on ollut tarpeellinen, niin ongelmiksi kuitenkin muodostui suostumuksen hallinnan keinot, joilla pyrittiin auttamaan yrityksiä dataoikeuksien käyttämisessä (Sitra, 2019). On epätoivoista odottaa, että yritysten toiminnan särmät hioutuvat aikojen saatossa tarpeeksi vain hallinnollisia sakkoja jakamalla. Siksi on tärkeää pyrkiä kehittelemään uusia konsepteja ja lähestymistapoja, joilla voitaisiin edistää reilua datataloutta. Tutkielmassa esiteltiin yksi innovatiivinen lähestymistapa henkilökohtaisen datan hallintaan, jossa digitaaliset oikeudet yhdistyvät kaupallisiin tarpeisiin: MyData-malli.

Valitettavaa on kuitenkin huomata, että vaikka datatalouden arvo ja mahdollinen potentiaali on ollut kovassa nousussa viime vuosina, niin Eurooppa on silti Yhdysvaltoja ja Kiinaa jäljessä, koska Euroopalla kerrotaan olevan vain 3% kaikkien alustayhtiöiden markkina-arvosta, kun taas Kiinalla luku on 30% ja Yhdysvalloilla omistus on 66% (Kubra Consult, 2017). Voisiko reilun datatalouden saavuttaminen olla avaintekijä Euroopan kilpailukyvyn vahvistumiseen? Tästä menestyksestä huolimatta Yhdysvaltoja ja Kiinaa ei kannata pitää tietosuojan roolimalleina. Luvussa 4. paljastui Yhdysvaltojen heikko data oikeuksien suoja, kun vertailu paljasti, että osalla osavaltioista ei ollut olemassa olevaa yksilön dataan liittyvää tietosuojalakia. Kiina taas ei ole harjoittanut vapaan markkinan mallia ja on edustanut täydellisesti digitaalista autoritarismia laajalla sensuurimuurillaan ja tekoälyyn perustuvalla sosiaalisella pisteytyksellään (Sitra, 2019).

Yritykset ympäri maailmaa ovat joutuneet muuttamaan datan käsittely toimintatapojaan GDPR:n valmistuessa, jotta kyseisen tietosuoja-asetuksen vaatimukset täyttyisivät, kuten Tietosuojavaltuutetun toimiston tarjoama Tietosuojavastaavia koskevat ohjeet antoivat odottaa (Tietosuojavaltuutetun toimisto, 2016). Virallisten tietosuoja viranomaisten valvoessa jopa vastahakoiset yritykset ovat joutuneet tekemään toimia hallinnollisten sakkojen uhalla (Wolff, 2019). Argumentti GDPR:n puolesta on kuitenkin yksinkertainen ja näin tukee yksilön dataoikeuksia; Jos haluat tarjota palveluasi eurooppalaisille kuluttajille, niin pelisäännöt ovat nämä.

Datan käytön sääntöjen ollessa selvät voidaan keskittyä reilun datatalouden visioimiseen. Päätavoitteena tähän on yksilöiden hallinta oman datansa yllä, jota GDPR ja MyData-liike tukee periaatteillaan (EUR-lex,2016; Poikola et al., 2018). Datasta saatiin arvoa vain sitä hyödynnettäessä ja se ei kulunut, kun sitä käytettiin muiden resurssien tapaan (Poikola et al., 2018). Näin on vaikea rajoittaa, että kuinka moni osapuoli käyttää dataa ja on helpompi ymmärtää datan myynnin syitä. Ihmisten yksityisyyden ja datan perusteella syntyvien innovaatioiden välille tulee rakentaa soviteltu ratkaisu, joten olisiko yleistynyt MyData-malli sopiva järjestely tälle vastakkainasettelulle?

Vaikka tutkielma laajasti kritisoi ja voi maalata henkilökohtaisen datan hyödyntämisen pahana, niin on kuitenkin hyvä muistaa sen monet tuomat hyödyt. GDPR:n tuomat datan käytön selvitys oikeudet auttavat yksilöitä ja tietosuojavaltuutettuja tunnistamaan, että mihin yksilön tietoja käytetään ja näin suurtenkin yritysten datan käyttö muuttuu läpinäkyvämmäksi yksilöille kuin ennen. Läpinäkyvyys auttaa varsinkin epäeettisen tiedonlouhinnan tai käsittelyn tapauksissa, kuten luku 3.3.1 osoitti Cambridge Analytican datan käyttötarkoitukset. Mutta tällaisissa tapauksessa ei tulisi tuomita datan louhintaa tai sosiaalista mediaa pahaksi, vaan sen epäeettiset periaatteet toiminnan taustalla.

6. Yhteenveto

Tämän tutkielman tarkoituksen oli tuoda esille yksilön dataan sisältyviä oikeuksia, GDPR:n vaikutusta ympäristössämme ja sekä pohtia reilun datatalouden määritelmää. Tutkielman perusteella todettiin, että EU on edelläkävijä yksilön dataan sisältyvissä oikeuksissa laajalla GDPR tietosuoja-asetuksellaan verrattuna muualle maailmaa. GDPR:n ulottuessa muihinkin maanosiin, yritykset ovat voineet varautua GDPR:n asettamiin käyttäjän tietoa koskeviin asetuksiin, mikä voi olla tulevaisuudessa hyvä asia, kun yksityisyyttä koskevat tietosuojalait yleistyvät. Tutkimukseen enemmän syventyessä kuitenkin paljastui, että GDPR ei tällaisenaan vielä täysin auta luomaan reilua datataloutta, vaan jatkuvia kehityksiä ja uusia innovaatioita tarvitaan samalla, kun teknologian mahdollistamat ilmiöt kasvavat.

Tämä kandidaatin tutkielma toteutettiin kirjallisuuskatsauksena yksilön dataan liittyvistä oikeuksista. Kirjallisuuskatsaus tarjosi paljon sisältöä näistä kyseisistä aiheista ja tiedon haku oli luontevaa. Tutkielma myös pyrki vastaamaan tutkimuskysymykseen: ”Miten voimme turvata yksilön dataoikeudet tulevaisuudessa?”. Tutkimuskysymyksellä pyrittiin rajaamaan tutkimuksen aluetta yksilön ympärille, mutta tutkimuksen edetessä yksilön näkökulman kirjoittamisen sijaan tutkielma päätyi kertomaan, että miten asiat reilussa datataloudessa menisi. Tähän oli syynä oivallus, että reilussa datataloudessa yksilöllä on hallinta dataansa ja näin myös yritykset pelaavat reiluilla yhteisillä pelisäännöillä tuottaen hyötyä yksilölle.

Kyseinen aihe myös tarjoaa mahdollisia jatkotutkimuksia. On erittäin todennäköistä, että tulevina aikoina yksilön dataan liittyviä oikeuksia ja tietosuojalakeja tullaan parantamaan maailmanlaajuisesti ja näiden vertailu GDPR:n välillä olisi tuonut lisää syvyyttä tutkielmaan. Alkuperäisenä ideana tutkielmassa oli ottaa vertailuun mukaan myös Kiina, mutta vertailu rajattiin Euroopan ja Yhdysvaltojen välille (ks. 4), kun Kiinan yksityisyyttä suojaavista laista löydetty tieto oli ristiriitaista ja vähäistä.

Lähdeluettelo

Ackoff, R. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, *16*, 3-9.

Apache Hadoop. (Ei päiväystä). Lainattu 15.6.2020, saatavilla: https://hadoop.apache.org/

Bellinger, G., Castro, D., & Mills, A. (2004). Data, information, knowledge, and wisdom. Lainattu 16.8.2020, saatavilla: https://homepages.dcc.ufmg.br/~amendes/SistemasInformacaoTP/TextosBasicos/Data-Information-Knowledge.pdf

Binns, R., Lyngs, U., Van Kleek, M., Zhao, J., Libert, T., & Shadbolt, N. (2018, May). Third party tracking in the mobile ecosystem. In *Proceedings of the 10th ACM Conference on Web Science* (pp. 23-31).

Blondel, V. (2013) MIT Technology Review: Data Sources. Lainattu 3.6.2020, saatavilla: https://www.technologyreview.com/2013/04/23/83721/data-sources/

Bruns, A., & Stieglitz, S. (2014). Twitter data: what do they represent? *it-Information Technology: Methods and Applications of Informatics and Information Technology*, *56*(5), 240-245.

Burgess, M. (2017) What is GDPR? The summary guide to GDPR compliance in the UK. Wired.co.uk. Lainattu 7.6.2020, saatavilla: https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018

Casado, R., & Younas, M. (2015). Emerging trends and technologies in big data processing. *Concurrency and Computation: Practice and Experience*, *27*(8), 2078-2091.

California Consumer Privacy Act of 2018. (2018). Senate bill No. 1121. Lainattu 16.8.2020, saatavilla: https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1121

Dwoskin, E. (2018). Facebooks bans Trump campaign's data analytics firm for taking user data. Washingtonpost.com. Lainattu 16.8.2020, saatavilla: https://www.washingtonpost.com/news/the-switch/wp/2018/03/16/facebook-bans-trump-campaigns-data-analytics-firm-for-taking-user-data/

EUR-lex. (2016). EU:n Yleinen tietosuoja-asetus. 4, 42 & 94 artiklat. Lainattu 16.8.2020, saatavilla: https://eur-lex.europa.eu/legal-content/FI/TXT/HTML/?uri=CELEX:32016R0679&from=FI

Finlex. (2018). Tietosuojalaki 1050/2018. Lainattu 16.8.2020, saatavilla: https://www.finlex.fi/fi/laki/alkup/2018/20181050

Finlex. (1999). Suomen perustuslaki. Lainattu 16.8.2020, saatavilla: https://www.finlex.fi/fi/laki/ajantasa/1999/19990731

Future of Privacy forum. (Ei päiväystä). Comparing privacy laws: GDPR v. CCPA. Lainattu 13.6.2020, saatavilla: https://fpf.org/wp-content/uploads/2018/11/GDPR_CCPA_Comparison-Guide.pdf

GDPR Enforcement Tracker. (2020). Lainattu 10.6.20 20, saatavilla: https://www.enforcementtracker.com/?

Hussey, P. (2020). Why Internet & Information Privacy Policy Is Necessary. *World Libraries*, *24*(1).

Hyry, J. (2019) Digitaalisten palveluiden käyttö- kysely. Lainattu 16.8.2020, saatavilla: https://media.sitra.fi/2019/01/16140605/digitaaliset-palvelut-kyselytutkimus-suomessa.pdf

IAPP. (2018). IAPP-EY Annual Privacy Governance Report 2018. Sivu 65. Lainattu 16.8.2020, saatavilla: iapp.org/media/pdf/resource_center/IAPP_EY_Gov_Report_2018.pdf

Import.io, (2019). What is Data Visualization and why is it important? Lainattu 15.8.2020, saatavilla: https://www.import.io/post/what-is-data-visualization/

Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, *51*(8), 56-59.

Jehl, L & Friel, A. (2018) "CCPA and GDPR Comparison chart". Lainattu 15.6.2020, saatavilla: iapp.org/media/pdf/resource_center/CCPA_GDPR_Chart_PracticalLaw_2019.pdf

Karhula, P. (Ei päiväystä). Tiellä sananvapauteen Digitaalisia jalanjälkiä seurattiin. Lainattu 3.6.2020, saatavilla: https://sananvapauteen.fi/artikkeli/2307

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences of the United States of America, 110(15), 5802–5805.

Kubra Consult. (2017). Digital Business Models and Platform Economy. Lainattu 14.6.2020, saatavilla: https://kubraconsult.blog/2017/11/04/digital-business-models-and-platform-economy/

Kuo, Y. H., & Kusiak, A. (2019). From data to big data in production research: the past and future trends. *International Journal of Production Research*, *57*(15-16), 4828-4853.

Laguna, R. (2014). Armed and Ready: How Your Data Is Being 'Weaponized' Against You. Lainattu 28.5.2020, saatavilla: www.wired.com/insights/2014/08/armed-ready-data-weaponized/

Lamdan, S. (2015). Social media privacy: A rallying cry to librarians. *The Library Quarterly*, *85*(3), 261–277.

Liikenne- ja viestintäministeriö. (2018). Suomi toimii omadata-mallin suunnannäyttäjänä. Lainattu 14.6.2020, saatavilla: https://www.lvm.fi/-/suomi-toimii-omadata-mallin-suunnannayttajana-980281

Manyika, J. (2011). Big data: The next frontier for innovation, competition, and productivity. Lainattu 10.7.2020, saatavilla: http://www.mckinsey.com/Insights/MGI/Research/Technology_and_Innovation/Big_data_The_next_frontier_for_innovation.

Matz, S. C., Kosinski, M., Nave, G., & Stillwell, D. J. (2017). Psychological targeting as an effective approach to digital mass persuasion. *Proceedings of the national academy of sciences*.

Murugesan, S., & Bojanova, I. (2016). Encyclopedia of cloud computing. John Wiley & Sons. Sivut 4-5.

Mydata.org. (Ei päiväystä) Our mission. Lainattu 16.8.2020, saatavilla: https://mydata.org/

Noordyke, M. (Ei päiväystä). US State Comprehensive Privacy Law Comparison. iapp.org. Lainattu 13.6.2020, saatavilla: https://iapp.org/resources/article/state-comparison-table/

Oikeusministeriö. (2018): Uusi tietosuojalaki voimaan vuoden 2019 alusta, lainattu 9.6.2020, saatavilla: https://oikeusministerio.fi/artikkeli/-/asset_publisher/uusi-tietosuojalaki-voimaan-vuoden-2019-alusta

Paycom. (2018). How employers can look to California as a trendsetter. Lainattu 15.6.2020, saatavilla: https://www.paycom.com/resources/blog/how-employers-can-look-to-california-as-a-trendsetter/

Pecánek, M. (2020). How to Use Google Analytics to Improve SEO Performance. ahrefs.com. Lainattu 15.8.2020, saatavilla: https://ahrefs.com/blog/google-analytics-for-seo/

Pitkänen, Olli. (2014). Sinun tietosi eivät ole sinun: rekisteröidyn oikeus hyödyntää omia henkilötietojaan. Oikeus. 43. 202.

Poikola, A., Kuikkaniemi, K., Kuittinen, O., & Honko, H. (2018). MyData–johdatus ihmiskeskeiseen henkilötiedon hyödyntämiseen.

Rouse, M. (2018). Data Modeling. lainattu 1.6.2020, saatavilla: https://searchdatamanagement.techtarget.com/definition/data-modeling

Salo, I. (2013). Big data: Tiedon vallankumous. Jyväskylä: Docendo.

SAS. (Ei päiväystä). Big Data. Lainattu 10.8.2020, saatavilla: https://www.sas.com/en_us/insights/big-data/what-is-big-data.html

SINTEF. (2013). Big Data, for better or worse: 90% of world's data generated over last two years. ScienceDaily. Lainattu 1.6.2020, saatavilla: www.sciencedaily.com/releases/2013/05/130522085217.htm

Sitra. (2019). Reilun datatalouden tiekartta. Lainattu 4.6.2020, saatavilla: https://www.sitra.fi/julkaisut/reilun-datatalouden-tiekartta/#1-johdanto-euroopasta-maailman-johtava-datatalous

Soken-Huberty, E. (Ei päiväystä). 10 Reasons why privacy rights are important. Lainattu 7.8.2020, saatavilla: https://www.humanrightscareers.com/issues/reasons-why-privacy-rights-are-important/

The Economist. (2017) "The world's most valuable resource is no longer oil, but data" Lainattu 28.5.2020, saatavilla: https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data

Tietosuojavaltuutetun toimisto. (ei päiväystä a). Tunne oikeutesi. Saatavilla: https://tietosuoja.fi/tunne-oikeutesi

Tietosuojavaltuutetun toimisto. (ei päiväystä b). Kun haluat tarkastella tietojasi. Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/kun-haluat-tarkastaa-tietosi

Tietosuojavaltuutetun toimisto. (ei päiväystä c). Mikä on henkilötieto? Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/mika-on-henkilotieto

Tietosuojavaltuutetun toimisto. (ei päiväystä d). Pseudonymisoidut ja anonymisoidut tiedot. Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/pseudonymisointi-anonymisointi

Tietosuojavaltuutetun toimisto. (ei päiväystä e). Erityisten henkilötietoryhmien käsittely. Lainattu 15.6.2020, saatavilla https://tietosuoja.fi/erityisten-henkilotietoryhmien-kasittely

Tietosuojavaltuutetun toimisto. (ei päiväystä f). Tietosuojavaltuutetun toimisto valvoo tietosuojaoikeuksiasi. Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/tietosuojavaltuutetun-toimisto

Tietosuojavaltuutetun toimisto. (ei päiväystä g). Tietoturvaloukkaukset. Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/tietoturvaloukkaukset

Tietosuojavaltuutetun toimisto. (ei päiväystä h). Henkilötietojen käsittelijät. Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/henkilotietojen-kasittelijat

Tietosuojavaltuutetun toimisto. (ei päiväystä i). Tietosuojavastaavat. Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/tietosuojavastaavat

Tietosuojavaltuutetun toimisto. (2016). Tietosuojavastaavia koskevat ohjeet. Lainattu 27.6.2020, saatavilla: https://tietosuoja.fi/documents/6927448/8316711/Tietosuojavastaavia+koskevat+ohjeet+fi.pdf/3aad84e5-bb59-4e64-bdaf-adc1e5f2d719/Tietosuojavastaavia+koskevat+ohjeet+fi.pdf/Tietosuojavastaavia+koskevat+ohjeet+fi.pdf

Tietosuojavaltuutetun toimisto. (2020). Tietosuojavaltuutetun toimisto määräsi kolme seuraamusmaksua tietosuojarikkomuksista. Lainattu 15.6.2020, saatavilla: https://tietosuoja.fi/-/tietosuojavaltuutetun-toimiston-seuraamuskollegio-maarasi-kolme-seuraamusmaksua-tietosuojarikkomuksista?fbclid=IwAR1Dp9hd5uPiyDSL-F2tKjuP3qy6j2JXgJhkYsC8eY6EdoKGDtBR7YzoI14

TNS Opinion & Social (2018). Europeans' Attitudes towards Cyber Security, sivut 35, 49. Lainattu 4.6.2020, saatavilla: https://op.europa.eu/en/publication-detail/-

/publication/388ac375-c438-11e7-9b01-01aa75ed71a1/language-en/format-PDF/source-50238852

Uusiteknologia.fi. (2018). MyData2018 kokoaa digitiedon soveltajat Helsinkiin. Lainattu 8.6.2020, saatavilla: https://www.uusiteknologia.fi/2018/07/25/mydata2018-kokoaa-digitiedon-soveltajat-helsinkiin/

Vastuugroup.fi. (ei päiväystä). Sinullahan ei ole mitään salattava? Lainattu 14.6.2020, saatavilla; https://www.vastuugroup.fi/fi-fi/mydatashare

Vice.com. (2017). The data that turned the world upside down. Lainattu 11.6.2020, saatavilla: https://www.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win

Vänskä, R., & Härkönen, T. (2020) Henkilödatan jäljillä. Lainattu 15.8.2020, saatavilla: https://media.sitra.fi/2020/06/22153308/henkilodatan-jaljilla.pdf

Wallace, D. P. (2007). Knowledge Management-Historical and Cross-Disciplinary Themes, s.13-14.

Wolff, J. (2019) How is GDPR doing? Slate.com. Lainattu 7.8.2020, saatavilla: https://slate.com/technology/2019/03/gdpr-one-year-anniversary-breach-notification-fines.html

Wolford, B. (Ei päiväystä). What is GDPR, the EU's new data protection law? Lainattu 10.8.2020, saatavilla: https://gdpr.eu/what-is-gdpr/

Zhang, S., Zhang, C., & Yang, Q. (2003). Data preparation for data mining. *Applied artificial intelligence*, *17*(5-6), 375-381.